The Moral Dilution Effect: Irrelevant Information Influences Judgments of Moral Character

Cillian McHugh1 & Eric R. Igou1

1 University of Limerick

Author note

All procedures performed in studies involving human participants were approved by institutional research ethics committee and conducted in accordance with the Code of Professional Ethics of the Psychological Society of Ireland, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. All authors consented to the submission of this manuscript.

Correspondence concerning this article should be addressed to Cillian McHugh, University of Limerick, Limerick, Ireland, V94 T9PX. E-mail: cillian.mchugh@.ul.ie

Abstract

Across five studies we investigated the moral dilution effect

*Keywords:* keywords

*Word count:* TBC

The Moral Dilution Effect: Irrelevant Information Influences Judgments of Moral Character

# Test

this is a paper template

# Study 1 - Bad Characters

The aim of Study 1 is to test if the dilution effect exists in the moral domain. Participants were presented with descriptions of four characters, two descriptions will only contain diagnostic information (morally relevant information) and two will additionally contain non-diagnostic information (non morally relevant information) along with the diagnostic information. We hypothesize that moral perceptions of the diagnostic only descriptions will be more severe than for the descriptions that also contain non-diagnostic information.

## Study 1: Method

### Study 1: Participants and design.

Study 1 was a within-subjects design. The independent variable was condition with two levels, diagnostic information only (diagnostic), and non-diagnostic information additionally included (non-diagnostic). We used the same two dependent variables as in Pilot Study 1, the four item moral perception scale (MPS-4) which showed good reliability, = 0.83, and the single item moral perception measure MM-1.

A total sample of 851 (303 female, 526 male, 14 non-binary, 5 other; 3 prefer not to say, *M*age = 26.11, min = 18, max = 76, *SD* = 10.14) started the survey. Participants were recruited from the student population at University of [BLINDED].

Participants who failed both manipulation checks were removed (*n* = 100), leaving a total sample of 801 participants (283 female, 496 male, 14 non-binary, 5 other, 3 prefer not to say; *M*age = 26.25, min = 18, max = 76, *SD* = 10.20).

### Study 1: Procedure and materials.

As in the pilot study, data were collected using an online questionnaire presented with Qualtrics (www.qualtrics.com). Participants were presented with four descriptions of characters (*Sam*, *Alex*, *Francis*, *Robin* from Pilot Study 1). All descriptions included diagnostic information relating to three moral foundations, e.g., *Imagine a person named Robin. Throughout their life they have been known to physically hurt others, treat some people differently to others, and show lack of loyalty*. We programmed our survey to randomly present non-diagnostic information along with two of the descriptions participants read (this was done through blocking, for details on the blocks see full materials at <https://osf.io/mdnpv/?view_only=77883e3fbc3d45f1a35fe92d5318cb67>. This meant that all participants read two descriptions containing diagnostic information only, and two descriptions that additionally included non-diagnostic information. We hypothesized that the descriptions including non-diagnostic information would be rated as less severe than the diagnostic-only descriptions. Study 1 was pre-registered at <https://aspredicted.org/DVY_QN3>

## Study 1: Results

The means and standard deviations for MPS-4 for each scenario are as follows: *Sam*, *M*MPS-4 = 2.55, *SD*MPS-4 = 0.86, *Francis*, *M*MPS-4 = 3.05, *SD*MPS-4 = 0.97, *Alex*, *M*MPS-4 = 2.32, *SD*MPS-4 = 0.88, *Robin*, *M*MPS-4 = 2.13, *SD*MPS-4 = 0.91. There was significant variation depending on the description, *F*(3,2280) = 297.82, *p* < .001, partial 2 = 0.13. *Francis* appeared to be rated as more moral than each of the other characters (all *p*s < .001), while *Robin* was rated as less moral than each of the other characters (all *p*s < .001), while *Sam* was rated more favorably than *Alex* (*p* < .001).

The means and standard deviations for MM-1 for each scenario are as follows: *Sam*, *M*MM-1 = 23.94, *SD*MM-1 = 16.18; *Francis*, *M*MM-1 = 30.12, *SD*MM-1 = 17.86; *Alex*, *M*MM-1 = 20.55, *SD*MM-1 = 16.65; *Robin*, *M*MM-1 = 20.60, *SD*MM-1 = 17.06. There was significant variation depending on the description, *F*(3,2253) = 154.08, *p* < .001, partial 2 = 0.05. *Francis* was rated more favorably than all other characters (*p* < .001), *Sam* was the next most favorably rated character, rated significantly more favorably than both *Alex* and *Robin* (*p*s < .001), there was no difference between *Alex* and *Robin* (*p* = = .953).

We conducted a linear-mixed-effects model to test if condition influenced MPS-4 responses. Our outcome measure was MPS-4, our predictor variable was condition; we allowed intercepts and the effect of condition to vary across participants, and scenario was also included in the model. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, 2(8) = 816.91, *p* < .001. Condition significantly influenced responses to the MPS-4, *F*(1, 799.42) = 51.47, *p* < .001; and was a significant predictor in the model when controlling for scenario, = -0.08, *t*(799.42) = -7.17, *p* < .001, with the diagnostic descriptions being rated as more immoral than the non-diagnostic descriptions Figure 1.

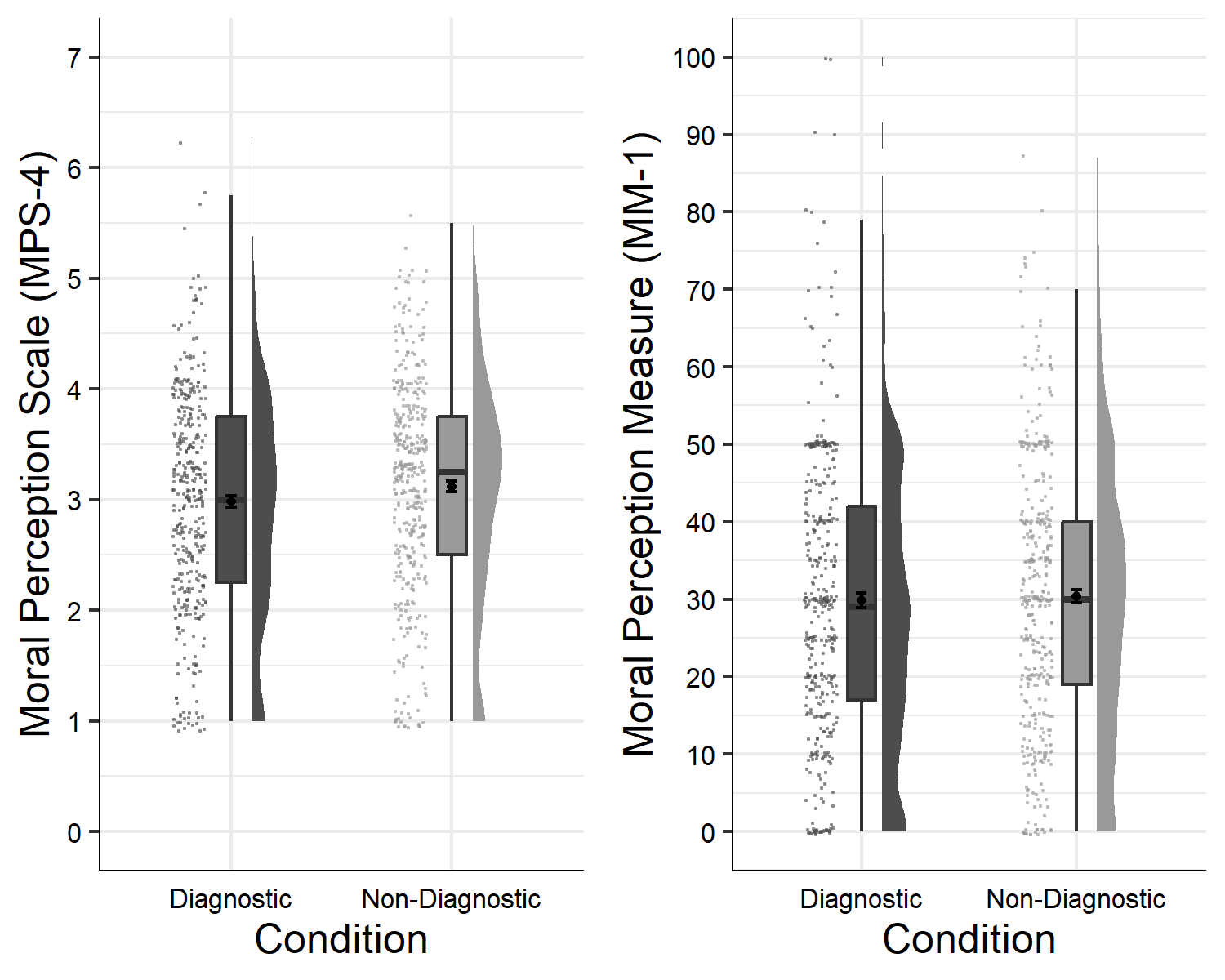


Figure 1: Study 1: Differences in moral perception depending on condition

We conducted a linear-mixed-effects model to test if condition influenced MM-1 responses. Our outcome measure was MM-1, our predictor variable was condition; we allowed intercepts and the effect of condition to vary across participants. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, 2(8) = 475.52, *p* < .001. Condition significantly predicted MM-1 responses *F*(1, 799.71) = 44.39, *p* < .001, and when controlling for scenario was a significant predictor in the model = -1.22, *t*(799.71) = -6.66, *p* < .001, with the diagnostic descriptions being rated as more immoral than the non-diagnostic descriptions Figure 1.

In the supplementary analyses we report the effect of condition on moral perception for each description individually.